

Recent Adaptions of the Devanagari Script for the Tibeto-Burman Languages of Nepal¹

Michael Noonan

University of Wisconsin-Milwaukee

1. INTRODUCTION: Most of Nepal's 120+ languages have no written tradition dating back more than a few years. When these previously unwritten languages come to be written for publication, which has been happening with some frequency since the promulgation of the new constitution in 1990, several practical problems arise, which include, but are not limited to, the following:

1. Which dialect or dialects should be used for writing?
2. Within the chosen dialect(s), which variant forms should be selected for writing?
3. Which script should be used?
4. How should the chosen variants be spelled given the conventions inherent in the script?
5. How should the script and its conventions be adapted to accommodate the phonological system of the language?

In this paper, I will be concerned primarily with the last two problems, specifically with respect to the 100+ Tibeto-Burman languages of Nepal and the attempts to write these languages with the Devanagari script. I will not attempt to survey all these languages; rather, I will report on six which have only recently come to be written with the Devanagari script: Chantyal, Gurung, Limbu, Sherpa, Tamang, and Thangmi.² These languages constitute a rough-and-ready sample of the Tibeto-Burman languages of Nepal without a tradition of writing in the Devanagari script³: some are locally prominent languages with large numbers of speakers, others more obscure; some have traditions of writing in scripts other than Devanagari, others do not. For none of these languages is there a fully standardized orthography, and none of these languages has been written in the Devanagari script in any serious way until quite recently.

The goal of this paper is to examine the ways in which the Devanagari script has, and has not, been adapted to the phonologies of the six languages. I'll first discuss the orthographic histories of each language. I will then discuss the characteristics of the Devanagari script and the phonological system it is designed to accommodate. Then I will present brief sketches of the phonological inventories of the six languages, follow-

¹ The work reported on in this paper has been supported by the following grants from the National Science Foundation: DBC-9121114, SBR-9600717, and SBR-9728369. I would like to thank Warren Glover, Tej Kansakar, Martine Mazaudon, and Mark Turin for assistance in writing this paper.

² Chantyal, Gurung, and Tamang are all members of the Tamangic group, a fairly straightforward grouping with a time depth more or less comparable to that of the Romance languages. Sherpa is a member of the Tibetan Complex, relatively closely aligned to the Tamangic group within the Bodish subphylum. Limbu is a Kiranti language. Thangmi, though typologically similar to the Kiranti languages, is not closely related to them: it's closest relative is Baraam, and the two may be related to the Western Himalayish languages, such as those in the Kanauri group.

³ Of the Tibeto-Burman languages of Nepal, only Newari has a long-standing tradition of writing using the Devanagari script.

ing this with a discussion of how the script has — and has not — been adapted to accommodate these sound systems. I will then discuss the issue of ‘superfluous graphemes’, and finish up with some concluding remarks on how Devanagari has been — and could be — adapted for these languages.

2. HISTORY OF WRITING IN DEVANAGARI: In this section, I’ll briefly discuss each of the six languages, sketching the history of writing in Devanagari for each and the controversies, where they have existed, over the choice of script. I will concentrate on publication rather than informal writing as might be found, for example, in personal correspondence. In any case, for most speakers of these languages, letter writing has until very recently been done exclusively through the medium of another language, usually Nepali, though Tibetan Buddhists might use Tibetan. The exclusive use of Nepali or Tibetan in literacy training coupled with the lack of established orthographic traditions for the languages of our sample has brought about and maintained this situation. With the exception of Limbu, it is only with the worldwide ethnic revival, a significant part of the Nepalese social and intellectual scene only since the 1970s, that writing in these languages has been considered a desideratum by significant numbers.

2.1 CHANTYAL: There were no publications in Chantyal until 1987, and indeed there have been no publications exclusively or even primarily in Chantyal save Bhulanja & Noonan (1995). That first publication in 1987 and the few publications which have appeared subsequently⁴ have all used the Devanagari script. Unlike all the other languages discussed in this paper, there has been no discussion about the use of any other script for this language. Bhulanja & Noonan proposed spelling conventions for Chantyal, appended to a collection of children’s stories that were distributed to schools in Chantyal-speaking villages. These conventions have not been observed in subsequent publications.

2.2 GURUNG: There have been advocates for four distinct scripts in recent years for Gurung (Glover 2002): the Tibetan script, particularly among Gurung Buddhists; the Khemaa lipi खेमाँ लिपि, an indigenous adaption of the Devanagari script; the roman script, advocated mostly by veterans of the British and Indian Gurkha regiments; and the Devanagari script. While the other scripts still have their adherents, Devanagari is clearly the script of choice for publication in Gurung.

There have been a significant number of publications in Gurung,⁵ at least by the standards of Nepal’s minority languages. Within the general conventions imposed by the Nepalese version of the Devanagari script, each author has felt free to invent his own spellings. Glover (2002) proposes a number of spelling conventions for Gurung publications, in particular those that might be used for a forthcoming Gurung-Nepali-English dictionary. A number of these proposals will be discussed below.

⁴ Chhantyal (2044 [1987], 2051 [1994]), Bhulanja & Noonan (1995), Nepal Chhantyal Shangha (2001), *Chhantyal* magazine.

⁵ For example, D. Gurung (2052 [1995]), P. Gurung (1997), R. Gurung (1998), Tamu (1996). The Glover, Glover & Gurung (1977) dictionary has a Nepali-Gurung index in Devanagari script.

2.3 LIMBU: Limbu has a script of its own, known as the *Sirijonga* script, which has been in use since at least the early 19th century. The script has undergone a number of revisions over the last century and is now used in schools in Sikkim, where Limbu is taught as a vernacular language subject in Limbu-speaking areas, and for publications in Limbu printed in Nepal. Despite the pride many Limbus feel for the Sirijonga script [which, incidentally, has a Unicode encoding], the Devanagari script is still used for Limbu due to the fact that Limbus are bilingual in Nepali and achieve familiarity with Devanagari at state schools which are run exclusively in the national language. Nonetheless, the Sirijonga script seems to be gaining ground. In any case, one consequence of the existence of an alternative script which is reasonably well adapted to the peculiarities of Limbu is that when the Devanagari script is used for Limbu, its deficiencies for representing Limbu phonology are evident to writers of the language. The Unicode now has coding for a glottal stop for the Devanagari script largely because of pressure from those who wished to use the Devanagari script for writing Limbu.

2.4 SHERPA: Until recently, Sherpa was written mostly in the Tibetan script, usually with Classical Tibetan spellings. Recent publications though are likely to be in the Devanagari script. There is a weekly newspaper, numerous teach-yourself Sherpa texts using Devanagari in varying degrees, and an online dictionary⁶, along with some literary publications. An informal standard has been emerging using more-or-less standard Nepali orthographic conventions and thus failing to fully accommodate the phonological structure of the language.

2.5 TAMANG: Tamang has come to be written only recently. Almost all publication has been in the Devanagari script⁷, though there are vocal advocates in the Tamang community for the Tibetan script, including a simplified version of the script referred to as *Tamyig*. Martine Mazaudon, at the invitation of the Tamang community, prepared a paper [Mazaudon 1993] discussing ways in which the Devanagari and Tibetan scripts could be adapted so as to accurately transcribe Tamang's phonemic distinctions. Recent publications do not appear to have made use of her suggestions, however.

2.6 THANGMI: As of this writing, there appears to have been only one publication of any sort in the Thangmi language: a Nepali-Thangmi-English dictionary prepared by Mark Turin and Bir Bahadur Thami [Turin & Thami 2003], using Devanagari script for the Thangmi entries, with conventions worked out by Turin in consultation with members of the Thangmi community.

3. THE DEVANAGARI SCRIPT: In order to discuss how the Devanagari script might be adapted for the six languages under consideration in this paper, we first must consider the characteristics of the script itself. Since the characteristics of the Devanagari script are well known and well documented, it will suffice to provide only a brief description which will introduce some terms that will figure later in the discussion.

⁶ <http://www.nepalresearch.org/>

⁷ The website <http://www.tamangghedung.org/> contains a listing of publications in Tamang.

William Bright (1996) has characterized the script aptly as an *alphasyllabary*, a useful term describing an orthographic system which has graphemes for consonants and vowels [like an alphabet] but writes consonant-vowel combinations as units [like a syllabary], in which the vowel graph acts as a *diacritic* to the consonant graph.⁸ When consonant graphs are written without a specific vowel graph, they are interpreted as representing a sequence of the consonant and an *inherent vowel*, often pronounced [ʌ] or [e]. The inherent vowel is canceled when a vowel diacritic is present or when a special diacritic known as the *virāma* is added to the consonant graph, indicating that no vowel follows the consonant. Consonant clusters often require special allographs, often for the first consonant, sometimes for the second as well; some consonant clusters are completely irregular and require special graphs of their own. These special allographs are referred to as *conjunct* or *combining* forms. Vowel nasalization is indicated by means of a diacritic known as the *candrabindu*; homorganic nasals can be transcribed with the *anusvāra*.

The free-standing allographs of the vowels and the consonants [the latter being interpreted as followed by the inherent vowel] are given in (1).

(1) THE DEVANAGARI SCRIPT

CONSONANTS	Labial	Den-Alv	Retroflex	Palatal	Velar	Glottal
Unaspirated Stop/affricate	प	त	ट	च	क	
Aspirated Stop/affricate	फ	थ	ठ	छ	ख	
Voiced Stop/affricate	ब	द	ड	ज	ग	
Murmured Stop/affricate	भ	ध	ढ	झ	घ	
Voiced Nasal Stop	म	न	ण	ञ	ङ	
Voiced Lateral Approximant		ल	ळ			
Voiced Tap Approximant		र				
Voiceless Fricative		स	ष	श		
Murmured voice [glottal fricative]						ह
Glide				य	व	
VOWELS:	अ, आ, इ, ई, उ, ऊ, ए, ओ					
DIPHTHONGS:	ऐ, औ					
SYLLABIC LIQUIDS:	ऋ, ॠ, लृ, लृ					
VISARGA <ḥ>:	:					
NASALIZATION:	◌̣, ◌̤					

The *candrabindu* [or *anunāsika*] <◌̣> is used to represent nasal vowels; the *anusvāra* <◌̤> is used to represent nasals homorganic to following consonants and sometimes nasal vowels. The *visarga* represents <ḥ>, but only syllable finally: its use is mostly restricted

⁸ It is not the syllable as such, but rather an orthographic unit known as the *akṣara* (Bright 1996), which constitutes a basic unit within this orthographic system.

to Sanskrit and Sanskrit borrowings.⁹ The retroflex lateral and the syllabic laterals are likewise restricted to Sanskrit and are exceedingly rare even there, but would likely be present in the typesetter's kit and therefore available to be called into service.

In (2) are summarized the phonological units that the script can – without additions, or modification and/or reinterpretation of the basic character set – represent phonemically. In the transcription used here, <^h> indicates aspiration, <ḥ> murmur.

(2) THE PHONOLOGICAL SYSTEM DIRECTLY REPRESENTED BY THE DEVANAGARI SCRIPT:

CONSONANTS:	Labial	Den-Alv	Retroflex	Palatal	Velar	Glottal
Unaspirated Stop/affricate	p	t	ʈ	c	k	
Aspirated Stop/affricate	p ^h	t ^h	ʈ ^h	c ^h	k ^h	
Voiced Stop/affricate	b	d	ɖ	j	g	
Murmured Stop/affricate	bḥ	dḥ	ɖḥ	jḥ	gḥ	
Voiced Nasal Stop	m	n	ɳ	ɲ	ŋ	
Voiced Lateral Approximant		l	ɭ			
Voiced Tap Approximant		r				
Voiceless Fricative		s	ʂ	ʃ		
Murmured voice [glottal fricative]						ḥ
Glide				y	w/v	

VOWELS:

i, i:, ĩ, ĩ:	u, u:, ũ, ũ:
e, ě	o, õ
	Λ, ̃Λ
	a, ̃a

DIPHTHONGS:

Λy, ̃Λy	Λw, ̃Λw
---------	---------

[other diphthongs can be written by juxtaposing vowel graphs]

SYLLABIC LIQUIDS:

ɾ, ɾ:	l, l:
-------	-------

The character <ṛ> has a murmured pronunciation in Nepali, which is what is indicated in the chart. The *visarga* <ḥ>, representing [h], is not listed since its use is restricted to syllable-final position.

For some languages using the script, in particular Hindi, a subscript dot is used to indicate certain consonant sounds not otherwise represented by the script. These are listed in (3):

- (3) क [qΛ], ख [xΛ], ग [ɣΛ], ज [zΛ], ङ [tΛ], ढ [ʈḥΛ], फ [fΛ]

Finally, a few words should be said about the Devanagari script as used for writing Nepali. Nepali is the national language of Nepal and the language of public education

⁹ A few orthographies have put the *visarga* to other uses, e.g. Newari, which uses it to indicate vowel length. Note that in Sanskrit, the *visarga* represents voiceless [h], not murmured [ḥ], and thus the visarga is not an allograph of <h>, which is always murmured.

and of interethnic discourse. All Nepalis below the age of, say, 60 are effectively bilingual in Nepali, and to the extent that they are literate, they achieve literacy first in Nepali. As we will see, these facts have a considerable influence on the way that the Devanagari script has been used for the six languages under consideration here.

The sound system of Nepali is represented quite well by the Devanagari script. A few graphemes are superfluous in the sense that they represent potential phonological distinctions which are not made – or, in some cases, not normally made – when speaking Nepali. These include:

The distinction between the sibilant characters:

- (4) श, ष, स *potentially* [ʃ], [ʂ], [s]

The sounds represented by these characters are occasionally distinguished by educated Nepalis, but are not ordinarily distinguished in casual speech.

The distinction between long and short <i> and <u>:

- (5) इ, ई *and* उ, ऊ

The syllabic liquids and the retroflex lateral [the last found only in forms from Vedic Sanskrit]:

- (6) ऋ, ॠ, लृ, लृ, ॡ

The *anusvāra*, which is often used to write homorganic nasals:

- (7)

The *visarga*, which, in any case, is found only in Sanskrit borrowings:

- (8)

Many speakers of western dialects of Nepali fail to distinguish the dental and retroflex series, though these are distinguished in Standard Nepali and in the Nepali taught in the schools; the labiodental fricative and the bilabial stop are also not regularly distinguished.

The only sounds found in Nepali that have no dedicated characters in the traditional Devanagari script are the retroflex flaps [ɾ] and [ɽ]. Ordinarily, these are written with the characters for the voiced and murmured retroflex stops, though in a very few publications, mostly of a pedagogical nature, they may be written with the subscript dot beneath the retroflex stop characters, as in Hindi:

- (9) ढ [ɽ] ढ [ɽ]

There is only one orthographic innovation found in Nepali *vis-à-vis* other languages using the Devanagari script. This involves the *repha*, the combinatory form of the <r> grapheme. Before <y>, the *repha* is often written as in (10),

- (10) र्य /ryΔ/

rather than as in (11), the standard method for writing the <ry> sequence word-medially in Devanagari:

- (11) र्य /ryΔ/

4. PHONOLOGICAL SYSTEMS OF THE SIX TIBETO-BURMAN LANGUAGES: Now we will discuss the six Tibeto-Burman languages that are the objects of this study. Below are charts

containing the phoneme sets of the six languages. In each chart, phonemes that would require unconventional representations in the Devanagari script, but which require no modification or reinterpretation of characters are placed within plain rectangles; phonemes which cannot be accommodated within the standard Devanagari script are placed within rectangles with double lines.

(12) CHANTYAL (Noonan 2003)

	Labial	Alveolar	Palatal	Velar	Glottal
CONSONANTS:					
Unaspirated Stop/affricate	p	t	c	k	
Aspirated Stop/affricate	p ^h	t ^h	c ^h	k ^h	
Voiced Stop/affricate	b	d	j	g	
Murmured Stop/affricate	b ^h	d ^h	j ^h	g ^h	
Murmured Stop/affricate with Voiceless Onset	p ^h	t ^h	c ^h	k ^h	
Murmured Stop/affricate with VI Asp Onset	th ^h		kh ^h		
Voiced Nasal Stop	m	n	ɲ	ŋ	
Murmured Nasal Stop	m ^h	n ^h	ɲ ^h		
Voiced Lateral Approximant		l			
Murmured Lateral Approximant		l ^h			
Voiced Tap Approximant		r			
Murmured Tap Approximant		r ^h			
Voiceless Fricative		s			
Murmured Fricative		s ^h			h
Glide			y	w	
Murmured Glide			hy	hw	
VOWELS:					
	i, ĩ,				
	e, ě				
		u, ũ,			
		o, õ			
		ʌ, ʌ̃			
		a, ă			

Murmur could be considered a tone or a property of syllables in Chantyal, but is represented here as a property of initial consonants (Noonan 2003). This corresponds to the intuitions of Chantyls, though in truth these intuitions may have been influenced by the conventions of the Devanagari script. In any case, in writing, Chantyls inevitably treat murmur as a property of the syllabic onset, as we shall see.

(13) GURUNG (Glover 1974, 2002)

	Labial	Den-Alv	Retroflex	Palatal	Velar	Glottal
Unaspirated Stop/affricate	p	t	ʈ	c	k	
Aspirated Stop/affricate	p ^h	t ^h	ʈ ^h	c ^h	k ^h	
Voiced Stop/affricate	b	d	ɖ	j	g	
Voiced Nasal Stop	m	n			ŋ	
Voiced Lateral Approximant		l				
Voiced Tap Approximant		r				
Voiceless Fricative		s, ʈ				
Murmured Fricative						h

Glide

y

w

VOWELS:

i, ĭ
e, eː, ě
a, ă
aː, ăː
u, ũ
o, oː, ȯ

TONES:

four tones; two are murmured.

As in Chantyal, murmur may be associated with voiceless onsets and with resonant onsets. The length contrast in vowels is marginal except for the low vowels.

(14) LIMBU (Michailovsky 2002)

CONSONANTS:

Unaspirated Stop/affricate

Aspirated Stop/affricate

Voiced Nasal Stop

Voiced Lateral Approximant

Voiced Tap Approximant

Voiceless Fricatives

Glides

Labial

p

p^h

m

Dental

t

t^h

n

l

r

s

Palatal

c

c^h

ny

y

Velar

k

k^h

ŋ

w

Glottal

ʔ

h

VOWELS:

i, iː, iː²
e, eː²
a, aː, aː²
u, uː, uː²
o, oː²
ɛ, ɛː, ɛː²
ɔ, ɔː, ɔː²

Voiced consonants occur as positional variants of the voiceless consonants. The glottalized vowels are phonologically distinct from a vowel + glottal stop sequence, though phonetically they may not differ.

(15) SHERPA (Kelly 2003)

CONSONANTS:

Unaspirated Stop/affricate

Aspirated Stop/affricate

Voiced Stop/affricate

Voiced Nasal Stop

Lateral Approximant

Tap Approximant

Voiceless Fricative

Murmured voice [glottal fricative]

Glide

Labial

p

p^h

b

m

Dent-Alv

t, ts

t^h, tsh

d

n

l, l̥

r, r̥

s

Retroflex

t̠

t̠^h

ɖ

Palatal

c

c^h

j

ɲ

f

y

Velar

k

k^h

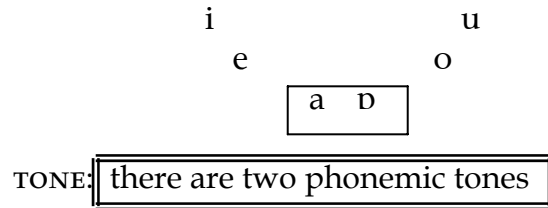
g

ŋ

w

Glottal

ɦ



Kelly (2003) maintains that there are two tones and a contrast, among the low tone words, between those with and those without murmured vowels: *e.g.* /gò/ ‘door’ *vs* /gò:/ ‘head’. This, no doubt, reflects an old tonal contrast, harking back to the time when Sherpa had a four-tone system, with two murmured tones.¹⁰

(16) TAMANG (Mazaudon 1993, 2003)

	Labial	Den-Alv	Retroflex	Palatal	Velar	Glottal
Unaspirated Stop/affricate	p	t	ʈ	c	k	
Aspirated Stop/affricate	p ^h	t ^h	ʈ ^h	c ^h	k ^h	
Voiced Nasal Stop	m	n			ŋ	
Voiced Lateral Approximant		l				
Voiced Tap Approximant		r				
Voiceless Fricative		s				
Murmured Fricative						ɦ
Glide				y	w	

VOWELS:

i, ɪ, i:, ɪ: u, ʊ, u:, ʊ:
e, ẽ, e:, ẽ: o, ɔ, o:, ɔ:
a, ɤ, a:, ɤ:

TONES: there are four tones; two are generally murmured

Phonetic voiced stops occur in Tamang intervocally, and also initially with the two low tones.

(17) THANGMI (Turin 2003)

	Labial	Den-Alv	Retroflex	Palatal	Velar	Glottal
CONSONANTS:						
Unaspirated Stop/affricate	p	t	ʈ	c	k	ʔ
Aspirated Stop/affricate	p ^h	t ^h	ʈ ^h	c ^h	k ^h	
Voiced Stop/affricate	b	d	ɖ	j	g	
Murmured Stop/affricate	bɦ	dɦ	ɖɦ	jɦ	gɦ	
Voiced Nasal Stop	m	n	ɳ		ŋ	
Voiced Lateral Approximant		l				

¹⁰ In any case, it would seem that there is a good deal of variation in the number of tonal contrasts in Sherpa: Kelly’s two [or three] contrast with the four reported in Hale (1973) and by Schöttelndreyer (1973). Hildebrandt (2003) and Noonan (ms) discuss the reduction or outright disappearance of tonal distinctions in the Bodish languages Manange and Nar-Phu under areal pressure from Nepali.

Voiced Tap Approximant	r			
Voiceless Fricative	s			
Murmured voice [glottal fricative]				ɦ
Glide		y	w	
VOWELS:				
	i		u	
	e		o	
		ʌ		
		a		

5. HOW THE SCRIPT HAS BEEN ADAPTED FOR THE SIX LANGUAGES: In this section I'll discuss the ways in which those phonological distinctions deemed problematic for the Devanagari script have been handled in published works in the six languages. Standardized orthographies utilizing the Devanagari script do not exist for any of the languages under consideration here, so I will report on the conventions used in published works that I have seen and on conventions proposed in works which discuss orthographic possibilities for individual languages.

I will not discuss orthographic conventions for diphthongs since the Devanagari script allows for the representation of diphthongs by the simple expedient of juxtaposing vowels: as long as the vowels can be written, the diphthongs they are found in can too.

In all, seven problematic phonological distinctions were uncovered in the last section: alveolar affricates, murmur, voiceless liquids, tone, vowel length, additional vowels, and the glottal stop. These will be discussed in turn.

5.1 ALVEOLAR AFFRICATES: Alveolar affricates contrast with alveopalatal affricates in Sherpa [and other languages in Nepal's Tibetosphere]. They are problematic orthographically only in the sense that they are considered unitary phonemes in these languages. In principle, there should be little difficulty writing the distinction in the Devanagari script: for example, त्स could be used for [tsʌ] and च for [tʃʌ]. In practice, the two series are written identically in all the samples of Sherpa that I've seen.¹¹

The reason for this no doubt stems from the fact that the members of the 'palatal' series in Nepali have two allophones for many speakers: an alveolar affricate and an alveopalatal affricate. Since Sherpas do not distinguish the two series in Nepali, they similarly fail to distinguish the two in Sherpa, even though the means to do so is easily available.

5.2 MURMUR: Apart from the cases where murmur is a concomitant of tone, murmur can be accommodated quite well within the Devanagari script. In Chantyal, where murmured syllables have voiceless and resonant initials, murmur is transcribed

¹¹ See, for example, the dictionary and other Sherpa materials available at:

<http://sherpa.nepalresearch.org/>

The romanization distinguishes the two series, but the Devanagari transcription does not.

with the character <ह>, which in Nepali represents [ɦ] not [h]. The syllable initial consonant is written in its conjunct form:

$$(18) \quad \begin{array}{lll} \text{पह} = /pɦ\Lambda/ & \text{थह} = /tɦ\Lambda/ & \text{मह} = /mɦ\Lambda/ \\ \text{सह} = /sɦ\Lambda/ & \text{चह} = /cɦ\Lambda/ & \text{कह} = /kɦ\Lambda/ \end{array}$$

In practice, Chantyls prefer to reverse the sequence of <ह> with syllable initial liquids and glides, producing:

$$(19) \quad \begin{array}{lll} \text{रह} = /rɦ\Lambda/ & \text{लह} = /lɦ\Lambda/^{12} & \text{वह} = /wɦ\Lambda/ \\ \text{यह} = /yɦ\Lambda/ & & \end{array}$$

This is just as well for <r>, as the sequence <rh> in initial position would be clumsy to affect in Devanagari. Either of the two possibilities in (20) would be required,

$$(20) \quad \begin{array}{ll} \text{रह} = /rɦ\Lambda/ & \text{रह} = /rɦ\Lambda/ \end{array}$$

and the second could, potentially, be misinterpreted as the sequence /rɦɦ\Lambda/ is possible in Chantyal.

5.3 VOICELESS LIQUIDS: Orthographically related to the previous problem is the problem of voiceless liquids. These issues are related because the usual means of representing voiceless liquids in the languages in which they are found is with <ह>, the same character used to represent murmur where there is no special character. In Nepali, <ह> is murmured, not voiceless; this would cause no particular problem in languages where <ह> was not also being used to represent murmur. The two languages in our sample that have voiceless liquids are tonal, and in both languages murmur is a concomitant of tone – of two of Gurung's four tones and of one [or two] of Sherpa's two [or three]. In either or both languages, therefore, <ह> could also be used to transcribe tone.

In Sherpa, <ह> is used to transcribe the voiceless tap and the voiceless lateral:

$$(21) \quad \begin{array}{ll} \text{ह} = /r̥a/ & \text{लह} = /l̥a/ \end{array}$$

Note that Sherpas use opposite orders of the liquid character and the <ह> character in writing the voiceless liquids. In the Sherpa documents I've seen, the <ह> character is also used with vowel-initial words with low tone: such words are pronounced with phonetic murmur.

In Gurung, Glover has proposed to render the voiceless lateral fricative with the sequence <kl>:

$$(22) \quad \text{क्ल} = /t̥a/$$

The justification for this is that the voiceless lateral fricative derived historically from the <kl> cluster, and some dialects still retain this cluster, which contains a prominent voiceless lateral [k̥l̥a].

5.4 TONE: Before going on to talk about how tone can be represented in the Devanagari script, I should say few words about how tone is manifested in the Tibeto-Burman languages of Nepal. Tone in these languages is not the domain of the syllable,

¹² I've also seen लह = /lɦɦ\Lambda/, which is easier to read.

as in the Sinitic languages, but rather of the word. Root morphemes have inherent tone, which spreads over the suffixes, which have no inherent tone – there are almost no prefixes in these languages. Therefore, in a perfectly phonemic orthographic system, it would only be necessary to specify tone once for each word: each syllable need not be marked.¹³

There is no straightforward way to write tones in the Devanagari script, and three of the languages under consideration here – Gurung, Sherpa, and Tamang – are tonal. However, since murmur is a concomitant of certain tones in all these languages, the Devanagari murmured stops and affricates and the character <ह> could be used to distinguish some of the tones from the others. This is in fact done in all these languages to some degree.

As noted, tones are not written in Sherpa save for the convention of writing <ह> in words with low tone beginning with vowels. This failure to mark tone extends also to minimal pairs reported by Kelly (2003) as being distinguished solely by murmur, for example in the online Sherpa dictionary referred to earlier. In this respect, Sherpa is alone among the languages in our sample in not using the Devanagari murmured stop and affricate characters to distinguish some tones.¹⁴

Tamang has four tones, two of which are typically accompanied by murmur or by voiced initial consonants. In works written by Tamangs, the two historically low tones¹⁵ are written with murmured or voiced initials; the murmured initials are written with either the existing Devanagari murmured stop and affricate characters or with sequences of consonants and <ह> as described above for Chantyal. The tonal distinctions between the two non-murmured tones and the two murmured tones are not written.

Mazaudon (1993) has proposed two techniques for writing all four of Tamang's tones, though I am not aware of publications which have taken up her suggestions. The first technique would employ numbers to mark tones in a manner analogous to the way tones are often indicated in linguist's transcriptions, *e.g.* publications by Mazaudon herself. The second would use voiced and murmured characters as noted above to indicate the two historic low tones, and use an apostrophe to distinguish the other tonal contrasts.¹⁶ This would result in tone transcriptions like the following, from Mazaudon (1993):¹⁷

(23)	PHONETIC TRANSCRIPTION	TO NE NUMBERS	MURMURED OR VOICED CHARACTERS + APOSTROPHE
'god'	¹ la	^१ ला	ला
'eye'	² mi:	^२ मी:	'मी:

¹³ For more information on this see Mazaudon (1973, 1978, 1993-4).

¹⁴ This, most likely, is further evidence for the decline of tonal distinctions in Sherpa noted earlier.

¹⁵ There is considerable variation in Tamang dialects as to how the contemporary tones are pronounced as Mazaudon has documented in a number of publications, for example Mazaudon (1993-4).

¹⁶ The use of the apostrophe in linguist's transcriptions of tonal contrasts in Tamang goes back at least to Hari (1970).

¹⁷ The transcription in (23) incorporates Mazaudon's suggestions for indicating vowel length. These will be discussed below.

'hill'	³ ri	री	रही
'thief'	⁴ jo	यो	यहो
'chest'	¹ ku	कु	कु
'chin'	² kam	काम	काम
'back'	³ ko	को	गो
'thatch'	⁴ ki	की	गी

An alternative possibility suggested by Mazaudon is that the subscript dot be used to indicate murmur, as in की 'thatch'. As noted, I'm not aware of any Tamang speakers using the superscript numbers, the apostrophe, or the subscript dot to indicate tone.

Gurung, like Tamang, has four tones. In writing these tones, Gurung writers have used the Devanagari murmured characters to indicate the two murmured tones, but most writers, again like the Tamangs, have not distinguished the two non-murmured tones and the two murmured ones. There is one exception, however. R.B. Gurung in his 1998 dictionary uses the murmured characters to write the murmured tones, and uses the Devanagari distinction between between long and short <i> and <u>, <इ> and <उ> versus <ई> and <ऊ>, to distinguish between the murmured and non-murmured tones when the roots contain either of these vowels. The long versions are used for what Sprigg (1995) refers to as 'contour tones', the short versions for the level tones. It turns out that the greatest functional load in terms of frequency in distinguishing the contoured and non-contoured pairs is on the <i> and <u> vowels, as noted by Glover (2002). So, while this device is not available for the other vowels [recall that Gurung /a/ and /a:/ are phonemically distinct], it is has some utility. [Recall also that there is no length distinction in high vowels in Gurung.] Examples, from Glover (2002), can be found in (24):

- (24)
- | | | | |
|---------------|--------------|---------------|-----------|
| NON-MURMURED | NON-MURMURED | MURMURED | MURMURED |
| NON-CONTOURED | CONTOURED | NON-CONTOURED | CONTOURED |
| कि | की | किह | कीह |

For the other vowels, only the murmured/non-murmured is marked:

- (25)
- | | | | |
|---------------|--------------|---------------|-----------|
| NON-MURMURED | NON-MURMURED | MURMURED | MURMURED |
| NON-CONTOURED | CONTOURED | NON-CONTOURED | CONTOURED |
| के | के | केह | केह |

5.5 VOWEL LENGTH: Three of the languages under consideration here have a phonemic distinction between long and short vowels: Gurung, Limbu, and Tamang. The Devanagari script can accommodate length distinctions in high and low vowels, but not in mid vowels, and all three languages distinguish long and short mid vowels. Limbu has, in addition, a quality distinction in mid vowels which will be discussed in the next section.

The length contrast in Gurung is found only in mid and low vowels and is quite marginal except in the low vowels, which can easily be accommodated in Devanagari.

Gurung authors do not write vowel length except in the low vowels, and write length for these vowels consistently.

Tamang distinguishes vowel length at all vowel positions. As noted, the Devanagari script is equipped with the means to write length distinctions with high and low vowels, though not with mid vowels. While this orthographic means is available for writing vowel length, it is not used consistently by Tamang writers. There are two reasons for this. First, Tamangs learn to use the Devanagari script in Nepali-medium schools, and in Nepali there is no phonemic difference in vowel length, though there is a quality difference between low vowels which is written with the <अ> and <आ> graphemes and their allographs. The orthographic long and short high vowel characters are simply a nuisance in learning to spell Nepali, since they are pronounced identically and are written on a non-transparent etymological basis. Tamang writers, therefore, do not associate these graphemes with the length distinctions in their language and write them in Tamang with no regard to their phonetic potential.

Second, with regard to the graphemes that can be used to write long and short low vowels: in Nepali these characters are used to write vowels which differ in quality, [ʌ]~[ɐ] *versus* [a], not in length, whereas in Tamang there is a true length distinction: [a] *versus* [a:]. Since Tamang has [a] but not [ʌ]~[ɐ], writers often use the long <आ> grapheme to write both long and short phonemes, particularly at the ends of words, though in other positions the the long and short vowels may be distinguished by the <अ> and <आ> graphemes, albeit inconsistently.

Needless to say, the length difference in mid vowels is not written at all.

Mazaudon (1993) has proposed that the Devanagari character *visarga* be used to indicate vowel length for Tamang. The visarga could be used either as the sole means for writing long vowels or as an adjunct to the Devanagari long and short vowel characters. If the visarga is used as the sole means, she suggests using the long vowel characters for the short vowels. Her proposals are illustrated in (26):

(26)	PHONETIC	VISARGA AS ADJUNCT	VISARGA AS SOLE MEANS
	ka	क	का
	ka:	का	काः
	ki	कि	की
	ki:	की	कीः
	ku	कु	कू
	ku:	कू	कूः
	ke	के	के
	ke:	केः	केः
	ko	को	को
	ko:	कोः	कोः

So far as I am aware, these conventions have not been observed by Tamang writers.


For Limbu, on the other hand, the visarga has been used for publications in the Devanagari script to indicate vowel length. Limbu writers are likely to know the Limbu script, in which the vowel length distinctions found in the language can be written efficiently. Since most writers are fully aware of the existence of these length distinctions and that they may be written, they are prepared to use available devices to write them when using the Devanagari script.

5.6 ADDITIONAL VOWELS QUALITIES: In our sample, only one language has distinctions in vowel quality that cannot be straightforwardly accommodated within the Devanagari script. Limbu distinguishes /ɛ/ from /e/ and /ɔ/ from /o/, as well as their long counterparts. The Limbu script has graphemes for these phonemes, and when Limbus write their language in the Devanagari script they are aware that the distinctions can be written. In Devanagari script, the lower mid vowels are written with the mid-vowel characters together with the subscript dot:

(27)	PHONETIC	LIMBU IN DEVANAGARI
	ke	के
	kɛ	के̎
	ko	को
	kɔ	को̎
	ke:	केः
	kɛ:	के̎ः
	ko:	कोः
	kɔ:	को̎ः

While it is certainly not the case that these conventions are always observed when writing the Devanagari script, it is not difficult to find printed examples utilizing them.¹⁸

5.7 GLOTTAL STOP: Two languages in our sample have a distinctive glottal stop, Limbu and Thangmi. In writing Limbu in the Devanagari script, a new glottal stop character, <?> has been come into use. As noted in section 2.3, there is now a Unicode encoding for this character within the Devanagari character set. In the printed samples I've seen, the roman interrogation sign <?> is normally used in place of this Devanagari character since typesetters already have this sign for setting English. Some recent computer typefaces for Devanagari – the Annapurna typeface, for example, available from the Summer Institute of Linguistics – currently include this character, which, incidentally, is identical to the phonetic character for a glottal stop. The character may also be written with a horizontal headstroke, as in (28):

(28) 

The first published Thangmi dictionary [Turin & Thami 2003] uses the novel technique of employing the halant, ordinarily used to negate the inherent vowel, to in-

¹⁸ For example Kaila (2052 [1995]).

dicating the glottal stop. In this use, the halant is appended to vowel characters. In (29), we see the Limbu and Thangmi conventions for indicating the glottal stop:

(29)	PHONETIC	LIMBU IN DEVANAGARI	THANGMI IN DEVANAGARI
	koʔ	कोʔ	को

6. SUPERFLUOUS GRAPHEMES: Earlier in the paper I discussed the fact that the Devanagari script contains more graphemes than are strictly necessary for a perfect phoneme-to-grapheme representation of Nepali. This results from the fact that the standard spelling is largely etymological and the alphabet is identical to the one used for writing Sanskrit. Because literacy in Nepal almost always begins in a Nepali language context, those writing the six languages of our sample transcribe obvious Nepali borrowings with their standard Nepali spellings insofar as possible, even when the distinctions the spellings encode are not used in the writer's own language — or in Nepali. Since some of these languages, most especially Chantyal,¹⁹ have borrowed vast numbers of Nepali vocabulary items, Nepali spelling conventions will necessarily leak into whatever additional conventions are required for the language being written.

Even when writing words clearly not borrowed from Nepali, Nepali spelling conventions — and, indeed, the look of Nepali written on the page — affect the way our six languages are written. For example, four of the languages do not contrast long and short high vowels. Nonetheless, most writers of those languages make use of both the long and short graphemes in writing. In general, the distribution seems to be based on what can only be described as aesthetic considerations. This is most pronounced with regard to the <इ> and <ई> graphemes: the long vowel grapheme, with an arc extending leftward over the horizontal bar is written at the ends of words; the short vowel grapheme, with an arc extending rightward over the horizontal bar, is written elsewhere. An example, taken from Chantyal, is found in (30):

(30) चिहन्जी /chinji/ 'finished'

It is remarkable how consistent writers are in applying this particular convention in these languages.

7. CONCLUSIONS: Two sorts of conclusions can be drawn from the discussion above. The first concern changes to the Devanagari script which have typographic consequences; the second concern general conclusions regarding the modes of adaption of the Devanagari script for the languages in our sample.

With regard to the first, a few conclusions are summarized in (31):

(31)

1. With the exception of the character for the glottal stop, which has already been assigned Unicode encoding, there are no new characters requiring special encoding which are regularly used by writers in the languages of our sample.

¹⁹ In Noonan (1996) it is estimated that 74% of Chantyal vocabulary is borrowed from Nepali.

2. There are some character combinations, commonly used for some of the languages of our sample, which are not found in Nepali and other Indic languages using the Devanagari script. These include combinations of <ॠ> and other consonants, the use of the subscript dot to indicate vowel quality, and the use of the halant to indicate the glottal stop. All of these, however, can be accommodated within standard Devanagari digital typefaces.
3. If Tamangs were to adopt Mazaudon's suggestions for using superscript Devanagari numerals for transcribing tone, these numerals would require special encoding — assuming that such coding has not already been assigned.

Some general conclusions are listed in (32):

(32)

1. It's clear that spelling conventions that people are exposed to in formal settings — in school, in literacy programs, etc. — even if these are in the medium of another language, determine how people will use the resources of the script when applied to their own languages.
2. Most writers adopt a rough and ready approach to writing their native languages, missing many distinctions and adhering to Nepali conventions wherever possible; where the Devanagari script using Nepali orthographic conventions fail to render a phonemic distinction, the distinction is ignored.
3. Where the Devanagari script is adapted to the phonologies of these languages, it is only because of the existence of appropriate graphemes *and* an appropriate pronunciation of those graphemes in the Nepali of the schools. The uses to which the <ॠ> has been put are examples of this principle.
4. The proposals of foreign linguists are usually ignored by native writers. However, the works produced by linguists using reformed orthographies [high prestige dictionaries, translations, etc.] are often a source of community pride and the conventions used in such works may ultimately influence writing in these languages. Adoption of these proposals by community organizations could be decisive, particularly if they are used in active native-language literacy campaigns, but all this has not yet happened.
5. The existence of a widely-used alternative script better adapted to the language, as we find in Limbu, encourages the use of alternative conventions in the Devanagari script by native language writers.

References

- Bhulanja, Ram Prasad & Michael Noonan. 1995. *Kyaratimaye Sastarma*. UW-Milwaukee.
- Bright, William. 1996. 'The Devanagari script.' In Peter Daniels & William Bright, eds. *The World's Writing Systems*. New York: Oxford University Press. Pp384-90.
- Chhantyal, Dil Bahadur Gharabja. 2044 [1987]. *Chhantyal Jati: Ek Parichaya*. Kathmadaun: Su. Shri Jayanti Gharabja Chhantyal.
- Chhantyal, Dil Bahdur Gharabja. 2051 [1994]. *Chhantyal Bhasama Kurakani*. Kathmadaun: Su. Shri Jayanti Gharabja Chhantyal.

- Glover, Warren. 2002. 'Choosing a Gurung orthography for a new dictionary.' *Gipan: TU papers in linguistics*, 2 (May): 25-38.
- Glover, Warren & Jessie R. Glover & Deu Bahadur Gurung. 1977. *Gurung-Nepali-English Dictionary, with English-Gurung and Nepali-Gurung Indexes*. Pacific Linguistics, Series C, No. 51. Canberra: Research School of Pacific Studies, ANU.
- Gurung, Dillijung. 2052 [1995]. *Tamu (Gurung)/ Nepali/ English dictionary*. Thamel: Hotel Marshyandi.
- Gurung, Prahlad. 1997. *Tamu yumai: tamu jaatiharuko ek sããskritik jhalak [A cultural glimpse of Gurung people]*. Darjeeling: Candra Tula Prakashan.
- Gurung, Rem Bahadur. 1998. *Gurung dictionary: Tamu tañ phuñ chhyoe (Gurung-Nepali-English)*. Tamu Chhoñj Dhiñ.
- Hari, Maria. 1970. 'Guide to Tamang tone.' Kathmandu: Summer Institute of Linguistics. ms.
- Kaila, Bairagi. *Tongsing tokma mundhum* ["Tongsing Ritual"]. 2052 [1995]. Kathmandu: Royal Nepal Academy.
- Kelly, Barbara. 2003. 'A Grammar and glossary of the Sherpa language.' In Carol Genetti, ed. *Tibeto-Burman Languages of Nepal: Manange and Sherpa*. Canberra: Pacific Linguistics.
- Mazaudon, Martine. 1973. *Phonologie Tamang*. Paris: Société d'études linguistiques et anthropologiques de France.
- Mazaudon, Martine. 1978. *Tibeto-Burman Tonogenetics. Linguistics of the Tibeto-Burman Area* 3.2 [Monograph issue].
- Mazaudon, Martine. 1993. 'Writing Tamang. A brief note on alphabetisation, spelling and transcription in Tamang.' *Syomhendo* (Kathmandu, Nepal). July-September 1993: 15-26.
- Mazaudon, Martine. 1993-4. *Problèmes de comparatisme et de reconstruction dans quelques langues de la famille tibéto-birmane*. Thèse: Université de la Sorbonne nouvelle: Paris III.
- Michailovsky, Boyd. 2002. *Limbu-English Dictionary of the Mew Khola Dialect with English-Limbu Index*. Kathmandu: Mandala Book Point.
- Nepal Chhantyal Sangha. 2001. *Chhantyal Bhasaka Kehi Shabda Jhan*. Kadmadaun: Nepal Chhantyal Sangha.
- Noonan, Michael. 1996. 'The fall and rise and fall of the Chantyal language.' *Southwest Journal of Linguistics* 15/1-2:121-36.
- Noonan, Michael. 2003. 'Chantyal.' In Randy LaPolla and Graham Thurgood, eds. *The Sino-Tibetan Languages*. London: Routledge.
- Noonan, Michael. ms. *The Grammar of Nar-Phu*. UW-Milwaukee.
- Schöttelndreyer, Heiderose. 1973. 'A Guide to Sherpa tone.' Kirtipur (Nepal): Summer Institute of Linguistics. ms.
- Sprigg, R K. 1995. "A tonal analysis of Gurung, with separate systems for register and contour pitch features." *Bulletin of the School of Oriental and African Studies* 60:448-54.
- Subba, B.B. 1979. *Limbu Nepali English Dictionary*. Gangtok: Govt. of Sikkim, Directorate of Education.

- Tamu, Dillijang. 2053 [1996]. *Tamu (Gurung) Bhasa Sikaun: Let's Learn Tamu (Gurung) Language*. Kadmandu: Tek Bahadur Tamu.
- Turin, Mark & Bir Bahadur Thami. 2003. *Nepali-Thangmi-English Dictionary*. Kathmandu: Martin Chautari.